



The International Journal for
Translation & Interpreting
Research
trans-int.org

Sign language interpreter aptitude: The trials and tribulations of a longitudinal study

Christopher Stone

University of Wolverhampton, UK

Christopher.Stone@wlv.ac.uk

DOI: 10.12807/ti.109201.2017.a06

Abstract: This paper discusses the process of undertaking an exploratory longitudinal study of language learning and interpreter aptitude. It discusses the context of aptitude testing, the test selection for a test battery, the recruitment of subjects within the small-scale study (n=22) and the administration of that battery within the context of whether longitudinal studies are feasible with small cohorts of sign language interpreters. Sign languages continue to be languages of limited diffusion in Europe. Even with gradually increasing numbers of ‘hearing’ sign language users, typically those wishing to become sign language interpreters do not have high levels of sign language fluency prior to enrolling in sign language interpreter training. As such, these students need to gain fluency in sign language, whilst also beginning to engage in interpreter education and interpreting-skills development. To date there is little understanding of how best to screen sign language interpreter program applicants to ensure the effective use of resources, i.e. to educate those who will both learn sign language to C1 fluency (Pro-signs, 2016) during the BA and also be able to learn how to interpret. Longitudinal studies enable us to take a longer view of learning and the professionalisation of skills and knowledge. They do, however, require significant time and this in itself can prove to be an obstacle when university researchers are required to produce tangible research outputs for career goals such as promotion or tenure.

Keywords: interpreter aptitude, interpreter expertise, cognitive development

1. Introduction

In the United Kingdom the profession of sign language interpreting (SLI) has gradually evolved from friends and family of deaf people providing informal and unpaid language brokering (Scott-Gibson, 1991; Stone & Woll, 2008), to trained professionals working for public services and the private sector. Following the implementation of the Disability Discrimination Act in 1995 (now subsumed within the Equality Act 2010) there has been an increasing demand for SLI, as this legislation places an obligation on providers of goods and services to make a ‘reasonable adjustment’ to ensure deaf people and people with disabilities have equality of service provision. For sign-language-using deaf people this ‘adjustment’ would often be the provision of a sign language interpreter. Since the UK ratified the United Nations Convention of the Rights of People with Disabilities in July 2009 (Office for Disability

Issues HM Government [ODI], 2011), this further obliges the provision of professional sign language interpreters within all aspects of civil society.

For the estimated 50,000-80,000 members of the Deaf community (ODI, 2011) whose preferred language is British Sign Language (BSL), there are only 952 registered sign language interpreters BSL/English Interpreters (NRCPD, 2016), i.e. those who would be considered professional sign language interpreters (see Stone 2013 for further discussion on what constitutes a professional interpreter vis-à-vis the interpreting profession and interpreting agencies). Currently, demand outweighs supply, meaning that either deaf people are unable to access services and engage in civil society, or language brokers are brought in who have not met the national occupational standards for interpreting (CFA, 2012), i.e. non-professional language brokers. Although there is no statutory regulation of sign language interpreters, service providers could be deemed to have fallen foul of their obligation under the Equality Act 2010, and the UNCRPD if they use the services of a non-professional interpreter.

Often the education routes available to those wishing to become sign language interpreters still reflect earlier models of interpreter training provision. That is to say that most potential interpreters begin learning sign language in recreational evening classes with students of mixed abilities and interests. These learners slowly progress to a level of proficiency where they can enter universities offering full-time training courses, or register for part-time vocational training routes (Brien, Brown, & Collins, 2002). Currently, there is also the possibility of taking an entry level or foundation year to gain some fluency in sign language, which enables attending students to achieve a robust A2 level of fluency on the CEFR (Pro-Sign, 2016) before entry into the ‘interpreting degree’, which is ostensibly a language learning and interpreting learning degree. These students are then expected to achieve C1 fluency (Pro-Sign, 2016) and to learn how to interpret within the BA/MA programs.

Since the establishment of a standalone register for sign languages interpreters (now known as the NRCPD – Simpson, 2007), public bodies, including more recently the Office of the Deputy Prime Minister and the Welsh Assembly, have funded short-term apprenticeship programs enabling trainees with varying levels of British Sign Language (BSL) to improve their language skills, observe working interpreters and begin to undertake interpreting assignments. In 2006 the Welsh Assembly Government announced a commitment of £1.6 millions of funding for ‘BSL Futures’—the future British Sign Language apprentice interpreter scheme for Wales, which also received £1.1 millions of matched funding to train “at least 30” BSL-to-spoken-language interpreters. Even if that project produced twice the minimum target (i.e. 60 interpreters), that would still represent £45,000 per trainee. This represents a significant investment per trainee; with limited resources available to ensure access for sign language using deaf people it would be appropriate to ensure they are used to support trainees who have an optimum potential. This was one of the motivating factors to undertake the study.

Currently both short-term training and long-term education programs have no objective selection criteria targeting language learning, in contrast to other intensive language training programs, such as those within the Foreign Office or military, which often use the Modern Language Aptitude Test (Carroll & Sapon, 2002) or similar tests from other institutions (e.g. the Colleges of Oxford language aptitude test).

Similarly, students are recruited without any research-informed assessment of their suitability for interpreting from the perspectives of language learning, cognition, or social attitudes. Interview panels rely largely on their intuition to determine the selection criteria for admitting trainees. As a result, candidates progress at unpredictable rates with varying levels of success. Even though there is an emerging literature addressing aptitude (Russo, 2011), there is minimal or no systematic approach to understanding factors that predict success within education programs. It was hoped that this longitudinal study would be able to support programs in better understanding the profile of students who can be successful within that type of educational environment.

Here, we report on an exploratory longitudinal study using a battery of language, psychomotor and cognitive tests to compare working sign language interpreters with those in undergraduate interpreter training programs. The goal of the test battery was to identify whether there were underlying language and cognitive abilities that could be used as a foundation to bootstrap the learning of BSL, and also to see whether either language ability or cognitive ability was predictive of being able to continue in the interpreting specialism of the course and predict examination results.

The exploratory nature of the longitudinal study aimed to enable us to gain a clearer picture on the role played by the cognitive function, language aptitude and linguistic skills on entry to the program, during interpreter education, and their subsequent development in training. We also wanted to compare these against the performance of working interpreters to better understand whether these played a role in language learning, interpreting development and professionalisation of interpreting skills. Longitudinal studies enable us to take a longer view of learning and the professionalisation of skills and knowledge. If we were successful in choosing appropriate tests, then this could inform screening for potential success in current interpreter education programs. However, in the UK small numbers of interpreters are trained every year, so ensuring that data can be collected systematically with sufficient data to draw robust conclusions on can be challenging.

We will now describe the background to the study, the selection of tests for the longitudinal study and the data collection method. We will then describe the challenges experienced, including securing funding and ensuring appropriate data collection timeframes.

2. Background

There are several questions that we need to answer with respect to the performance of interpreter students, as compared with the general population. As these students principally self-select, are they a sub-set of high performing students, i.e. those who somehow have the 'x' factor and although not 'born interpreters' (Mackintosh, 1999) have all of the underlying language and cognitive abilities to succeed? Or are they within the normal range of language and cognitive function and then the interpreter education is able to facilitate learning, in which case we should be able to identify key moments of learning as students begin to achieve test scores in a similar range to those of working interpreters. This meant that certain base-line test scores were required to see whether the interpreter students fall within the normal range of test cognitive function.

At this point it is also worth noting that several of the sign language interpreter education programs in the UK at the time of the study were housed in Deaf Studies departments. Although students were undertaking the interpreting specialism, if they were not able to learn BSL and pass interpreting exams at the level/pass mark required for the program they could be rehoused into the Deaf Studies program and still earn a BA. The hope for this study was that there would be sufficient numbers of those both able to complete the interpreting specialism and those rehoused in the Deaf Studies BA programs to enable comparison between those that had learnt BSL but not completed all of the interpreting program and those that had gone on to complete it. This would then also support the identification of the required language and cognitive function for student BSL/English interpreters to succeed within undergraduate programs in the UK.

The testing enabled the tracking of development at different stages of student interpreter education. The goal was to identify when stages of development occurred and if student interpreters began, at some stage, to perform in similar ways to working interpreters. The selection of working interpreters is discussed in greater depth below as are the testing points for the student interpreters during their education.

The study was designed in a research centre alongside linguists and psychologists. Within the research centre, there were several tests that were being used, or had been developed in-house to use to test the populations the research centre was interested in (sign language using deaf people). It was desirable to use these tests, so that different populations being researched within the research centre could be compared, if of interest, at a later date. This was one of the factors that influenced the design of the test battery.

The research centre also had a database of the (anonymised) test scores of all the participants engaged in the studies being undertaken. This meant that, if of interest for future studies, participants could be contacted about participating in further research projects.

Many of the tests used targeted specific language abilities, cognitive functions and cognitive abilities. As an exploratory study, while accepting that the students' interest would wane if the testing period was too long, the test battery was still longer (over two hours) than one might want for screening interpreting program applicants. We wanted to administer tests throughout the program, which enabled us to benchmark entry-level performance and track any improvements throughout language and interpreting learning. We will now describe the test battery in further detail.

3. The creation of the test battery

As noted above, the goal of the longitudinal study was to begin to identify appropriate tests that could screen applicants for sign language interpreter education programs. Any screening battery needs to have clear differentiation between weaker and stronger participants while being easy to administer and not too time consuming. To be able to differentiate requires us to understand and identify the development of language expertise and the cognitive skills we believe are developed in interpreting education.

At the time of the study, although there was some literature available on the testing of interpreter aptitude for conference interpreting (see Russo, 2011 for a description) there was no literature on community interpreting aptitude and only one study on sign language interpreting (Gómez et al., 2007).

If we consider each step of the process, firstly we need to consider language learning. In a recent study by Sparks (2009) children who performed well on L1 reading tests, amongst others, showed stronger skills in an L2 learned later at high school. Not only does this indicate transfer of skills from L1 to L2 but it also indicates that stronger L1 skills, which can be measured by reading fluency, can predict L2 skills. This suggested that we measure the fluency of the students and working interpreters with a reading test that included an adult age range.

We also wanted to see if a generic language aptitude test for spoken languages would also predict sign language learning aptitude as this had also not been tested previously. Gómez et al. (2007), who examined the initial aptitudes of interpreting students in Spain learning Spanish Sign Language (LSE), reported that ability to accurately repeat pseudo-signs (non-existent signs that follow the formal rules for LSE sign formation) was the only significant factor in predicting successful language learning and interpreting. Accordingly we wanted to include a non-sign repetition task. This was a complex task, but it also enabled the measurement of phonological development.

We wanted an objective measure as well as the students' exam results to judge whether students' fluency had increased and, if so, to what extent. As we had an in-house test that we had developed, we used it to measure comprehension fluency via grammaticality judgements.

We then needed to consider interpreting potential, i.e. whether the students needed to have any specific cognitive abilities to learn and/or undertake the interpreting task. Gómez et al. (2007) found that perceptual-motor and cognitive abilities played a more significant role than personality traits, so this became our focus too. For this we need some baseline measurements to be able to judge whether psychological traits were both within the normal range and then whether these improved with interpreter training. Some of these tests were standard nonverbal IQ tests (working memory capacity and nonverbal abstract problem solving); others arose because of discussions with other researchers (Brooke Macnamara and Šárka Timarová discussed below).

We will now discuss the language learning aptitude test, followed by the other in-house and standard language tests used to assess English and BSL. Then the focus will be on the cognitive assessments which tested: working memory capacity, nonverbal abstract problem solving, processing speed and flexibility, perceptual speed, psychomotor speed, task switching speed and conflict management and divided attention.

3.1 Language learning

As undergraduate students in the UK, who were predominantly white British women schooled in the UK, the cohorts of students had had some exposure to second language learning at school. Three of the students had a modern (foreign) language qualification taken at aged 16 (GCSE). Four of the students had had exposure to BSL prior to age five. Eleven other students had had some BSL tuition before arrival at the university. The remaining six students had limited exposure to learning a second spoken language, i.e. the modality of their first acquired language (L1), which can be written as M1L1 following Chen Pichler & Koulidobrova (2011). Even though the group was of mixed language exposure, the exploratory nature of the study meant that we were interested to see whether these were factors that influenced language learning and interpreting learning.

Initially we were interested in determining whether the students were able to use a general language (M1L1) learning aptitude to bootstrap their learning of a language in a new modality (M2L2), i.e. a visual, rather than auditory modality as there had been no studies on this prior to our study (since our study there have been other researchers exploring this for recreational language learners, e.g. Ortega and Morgan (2015)).

Carroll (1967) identified four components of spoken language learning aptitude:

1. Phonetic coding ability – the ability to perceive and remember distinct sounds associated with symbols
2. Grammatical sensitivity – the ability to recognize the function of a lexical element in a sentence
3. Rote learning ability – the ability to learn and retain associations between words in a new language and their meaning in English
4. Inductive learning ability – the ability to infer or induce rules governing the structure of a language

This then formed the basis of the modern language aptitude test (MLAT – Carroll & Sapon, 2002). The MLAT has been used to assess a variety of populations including undergraduate populations. The test has face validity for the learning of modern languages within an educational environment and had predicted the success of language learning for these populations. The MLAT is comprised of five different tests:

1. Number learning
2. Phonetic scripts
3. Spelling clues
4. Words in sentences
5. Paired associates

These tests result in individual scores for each test and a sixth score for the total test score. This test was used in order to identify whether there was any interaction between spoken language learning ability and the learning of sign language, identified either by one of the subtests or by the total score. We were interested to see, not only if this predicted BSL learning, but also whether any one subtest was more predictive than any other for language learning and interpreting performance in the student groups.

As well as overall language aptitude, we also wanted an objective measure of both English and BSL aptitude. Some of the students had agreed to share their university transcripts with us to enable a comparison of their examination results in BSL, translation, consecutive interpreting and simultaneous interpreting. However, we also wanted to be able to objectively gauge their language fluency. The approach taken and tests used will now be described below.

3.2 Language tests

Several tests were identified that were either used or developed in-house to measure the level of fluency of English readers and BSL signers. The revised Kirklees version of the Vernon Walden reading test (Hedderly, 1992) was used to test the accuracy and reading age of the participants in our study. The test is designed to rate the reading accuracy of children aged 8 to adults as a proxy for fluency. The rationale for using this test was to identify whether the

fluency of sign language interpreting students in their L1, which for BSL/English interpreting students is English in the main, was at an advanced level and whether this alone predicted better language learning and interpreting performance. This also allowed for the comparison of the English fluency of the students at entry and exit and subsequent comparison with the fluency level of the working interpreters.

The in-house research team had developed a BSL grammaticality judgement task (BSLGJT) (Cormier et al., 2012), based on the work of Mayberry and Boudreault (2006) but adapted for BSL. This task required students to watch a series of BSL videos and judge whether a sentence was a 'good' BSL sentence or not. The computer-based task was administered at several points throughout the interpreter student education and comprised of different language features such as: simple sentences; negated sentences; agreement verbs; interrogative sentences; relative clauses; and mimetic language/classifier constructions (Quinto-Pozos & Mehta, 2010). The aim of the study was to identify the point at which different linguistic structures were learnt during the undergraduate program. This could also have enabled the identification of those students with a visual-gestalt sensitivity, which allowed them to bootstrap sign language learning, i.e. learning a new language in the visual modality.

The visual-gestalt sensitivity in our study was measured by using a non-sign repetition task. The non-sign repetition task we administered used the stimuli created by Orfanidou et al. (2009) to test whether our participants could replicate sign-like non-signs accurately. Orfanidou et al. used the phonological rules of BSL to create non-signs that either obeyed the phonotactic rules of BSL but produced a string of sign-like activity, or disobeyed the phonotactic rules by adding additional phonological matter within the syllable structure. This built upon the work of Cutler and Norris (1988) and enabled us to test whether the interpreter students were able to mimic the non-signs, and then use that perception and psychomotor ability as a foundation for their language learning. It could enable us to gauge the robustness of the phonological representation of BSL in the students at different stages in their learning and for working interpreters.

Alongside the testing of language fluency and language development we wanted to target specific cognitive abilities. These tests will now be described.

3.3 Testing different cognitive abilities

The psychologists in the in-house team emphasised the need to ensure that the participants were within the normal range of cognitive function. This required the testing of several base measures. For this we used the digit span and matrix reasoning tasks from the WASI (Wechsler, 1999).

The digit span has two parts: forwards digit recall, where participants are asked to recall digits in order of presentation; and, backwards digit recall, where participants are asked to recall digits in reverse order of presentation. The forward digit recall tests working memory capacity and the backwards digit recall tests working memory, attention, encoding and auditory processing. There was no expectation that there would be an improvement in working memory, as the expectation was that other cognitive mechanisms were used to interpret (Liu et al., 2004; Wang & Napier, 2013). However, the nature of the longitudinal study did give us an opportunity to test whether this was actually the case.

The matrix reasoning task is a visual-spatial problem solving task. The test comprises a book of pictures containing a series of figures, one of which

is blank. The participants then have to select an appropriate figure to fill the blank from the options given. These scores gave us a measure of nonverbal intellectual ability, specifically testing nonverbal abstract problem solving, and inductive reasoning. Again while we did not expect to see an improvement in the matrix reasoning score, we were open to the possibility that learning a sign language would improve visual-spatial problem solving and the longitudinal study enabled us to test this too.

We also initially used the Speed and Capacity of Language Processing Test (SCOLP, Baddeley et al., 1992) to see whether this had sufficient discriminatory power to differentiate between the students' improvement during their program and any differences between students and working interpreters. After the initial test battery administration, the results did not differentiate between any groups and in several cases showed anomalous results. It was felt that the age of the test and the vocabulary it targeted were not appropriate for the groups.

As we are aware, the nature of interpreting involves processing information. To test for one aspect of this cognitive skill we used the paced auditory serial addition task or PASAT (Gronwall, 1977). This required the participants to listen to a series of numbers and to add pairs of number in a series. As stated in the instructions, 'The important thing is that you add the number you just heard to the number you heard right before it' (Brainmetric, 2004). The problem was remembering the number you have heard rather than adding the sum of the previous pair to the new number. It is a measure of cognitive function that specifically assesses auditory information processing speed and flexibility, as well as calculation ability.

At face value, this task would not appear to be ecologically valid for sign language interpreting: the modality difference of the languages implies that the auditory signal does not interfere with the visual-gestural signal. We were, however, interested to see whether this working memory capacity test yielded any correlations with the exam results of the students.

We were also interested in the risk-taking traits of the students. Students who are willing to take risks often are more likely to interact using the language being learned (Bacon & Finnemann, 1990). For the risk task, we administered the Barratt impulsivity scaleⁱ (BIS – Patton et al., 1995). This is a questionnaire that the participants fill in with 30 questions. The participants mark their answer to each question according to a four-point scale: rarely/never, occasionally, often and almost always/always. We were interested to see whether the risk taking was a trait that all of the language learners had, whether it was only found in one of the groups (interpreting students or rehoused students), and whether this was a trait shown by working interpreters.

3.4 Further cognitive tests

A fruitful meeting with Brooke Macnamara and Šárka Timarováa at CIT in Puerto Rico in 2008 led to the addition of several tasks. We also hoped this would enable some comparison of data sets or at least results that could be compared between the studies. Macnamara, later published as Macnamara et al. (2011) and Macnamara & Conway (2016) suggested using a variety of tests specifically targeting perceptual speed, psychomotor speed, task switching speed and risk taking.

ⁱ <http://www.impulsivity.org/measurement/bis11>

For perceptual speed, a pattern comparison (Salthouse & Babcock, 1991) was administered. This requires the participant to make decisions about whether two line drawings are the same pattern or a different pattern. Participants are presented with a printed page of patterns and with two rows of 15 pairs of down the page. The participants are given thirty seconds to decide and write against each pair whether they are the same (S) or different (D). We were interested to see whether the pattern discrimination and perceptual speed interacted with the learning of a visually perceived language, and whether this changed over time.

For psychomotor speed, a series of trail tests were used where the participants have to join the circles either in numerical sequence, or alphabetical sequence. This is the Connections A: numbers and Connections A: letters (Salthouse et al., 2000) and the combined scores give a composite psychomotor speed score. Here again we were interested in the interaction between perception and hand-to-eye coordination. We were interested to see whether psychomotor speed improved with the learning of a manually produced language.

For task-switching speed, the Connections B: numbers-letters and Connections B: letters-numbers (Salthouse et al., 2000) was administered. In this task the participants have to join the circles either in alternating numerical and alphabetical sequence, or alternating alphabetical and numerical sequence. The combined scores give a composite task-switching speed. This was a task that we hypothesised should show improved results with interpreter training. As such we were hoping that this would differentiate between students' performance early and later in the program and between those students rehoused in Deaf Studies and those students completing the interpreting program.

Finally, Timarova, later published as Timarova et al. (2014), suggested the use of a classic flanker task. The flanker task is something that has been demonstrated to be better performed by bilinguals (Costa et al. 2009) with the argument that the choosing between two languages develops the general cognitive abilities of conflict resolution between two competing alternatives. Evidence has, however, been found to demonstrate that bimodal bilinguals, i.e. those with a spoken and a signed language, do not demonstrate this control (Emmorey et al., 2008a).

Both of these studies include participants with childhood exposure to both languages. The bimodal bilinguals in the Emmorey et al. study were all from the US. In the US deaf community it is very acceptable to code-blend (Emmorey et al., 2008b) (i.e. to blend language codes from different modes together), which could account for the lack of cognitive control. We were interested to see whether the later learning and professionalisation of language use as a bimodal bilingual, within the sociolinguistic environment of having to choose one language over another, would result in better control in the flanker task.

Now that we have described the nature of the tests we will discuss the test battery administration.

3.5 The test battery administration

The final battery included several tests that were administered at the beginning and end of the students' program, i.e. year one semester one and year three semester three:

1. The MLAT (Carrol & Sapon, 2002)

2. A digit span task (Wechsler, 1999)
3. The matrix reasoning task (Wechsler, 1999)
4. The Kirklees revised Vernon Walden English reading test
5. A non-sign repetition task (Orfanidou et al., 2009)

There were also a series of tasks that were completed year one semester one, year two semester one, year three semester one and year three semester two:

1. A BSLGJT (Cormier et al., 2012)
2. Patterns (Salthouse & Babcock, 1991)
3. Connections A and B (Salthouse et al., 2000)
4. The PASAT (Gronwall, 1977)
5. The Barratt's impulsivity scale (Patton et al., 1995)

Table 1: Details of test and retest administration

	Semester 1	Semester 3	Semester 5	Semester 6
MLAT	X			
English fluency	X			X
Non-sign repetition	X			X
Digit span	X			X
Matrix reasoning	X			X
BSL GJT	X	X	X	X
Patterns		X	X	X
Connections A and B		X	X	X
PASAT		X	X	X
BIS		X	X	X

One of the issues was trying to ensure that we could recruit and test students as early as possible in semester one, so that we could get a baseline measure before too much intensive BSL learning had occurred. During the first year this was harder to do as the students needed to be asked if they were willing to participate, whilst settling into a new university course. In subsequent years, the program administrative support were able to ensure this happened closer to the start of the semester. However, for the first cohort of participants the data collection was six weeks into the semester and this could mean that by the time the initial testing occurred, the initial window had passed and performance had already changed due to language teaching and other factors.

A second issue is that the measures we have for interpreting performance were the exam results of the students. Rather than being a sole measure of interpreting performance, these exams included both performance and analysis results. As such they were not indicative of interpreting performance alone.

We will now give further information on the participants in the study.

4. Participants

The tasks were administered during the British Sign Language/English interpreter students' interpreter education program. To gain an n of 22 required three cohorts of students, meaning that the data collection itself took place over five years. Initially we collected data from 12 participants in cohort one, 4 participants in cohort two and 13 participants in cohort three. Initially 29 participants were recruited, although only 22 continued through the four test-administration dates for their respective cohorts leaving a total of 22 participants. All 22 performed within the normal range in the general cognitive abilities tests (digit span and matrix reasoning).

We know from other studies (Macnamara et al., 2011; Macnamara & Conway, 2016; and Timarová et al., 2014) that small sample sizes are not uncommon when collecting data from interpreters and interpreting students. The issue then becomes that weaker relationships may not be detected or reliable and so a large effect size would be required. However, with the limited number of undergraduates in sign language interpreter education, there is little that can be done to increase the number of participants without taking a longer time (data collection for this study took five years). Some of the tests had a large number of data points per participant, thus mediating the small sample size.

Table 2: Participants

	N	Age	Testing
BSL/Interpreting students	22 (5 male)	Median 19 (range 19-25)	4 times (3 x 1 year apart 1 x 1 semester apart)
Working interpreters	14 (2 male)	Mean 43 (range 34-53)	1 time

Following Ericsson's (2000/01) notion of expertise, we decided to recruit working interpreters who had at least ten years' experience. As we wanted to ensure that we were comparing like with like we also required the working interpreters to have at least a bachelor's degree. This degree did not need to be in interpreting as this would have limited the pool of interpreters available to participate in the study. Several interpreters had degrees and then had undertaken a two-year university diploma in sign language interpreting.

Finally, the working interpreters needed to have achieved their full professional status, which for sign language interpreters in the UK was passing a qualification mapped against the national occupational standards. Interpreters were recruited via email from the list of registered sign language interpreters (RSLI) with the NRCPPD. Similarly, all performed within the normal range in the general cognitive abilities tests (digit span and matrix reasoning).

5. Results

The MLAT was proven to have some validity in being predictive of language learning, which also appears to have some interaction with translation and interpreting.

The MLAT subtest one (number learning – phonological working memory) and BSL:

BSL exam semester one $r=.571$, $p=.006$

BSL exam semester two $r=.516$, $p=.017$

BSL exam semester four $r=.478$, $p=.033$

This makes sense in that BSL has no written form and so remembering biological strings of movement that must be remembered as linguistic matter is akin to the number learning task.

The MLAT total score also has a correlation with BSL exam semester one ($r=.492$ $p=0.032$), which suggests that other language learning factors are also in play during the first semester, i.e. not just phonological working memory. There is little interaction with the translation and interpreting modules but the advanced BSL (semester five), Translation (semester three), Consecutive Interpreting (semester four), Simultaneous Interpreting (semester five) and Interpreting Performance (semester six) are all correlated well with each other. BSL (semester one) does not correlate as well and so this would suggest more than initial language performance that contributes to interpreting performance. Fortunately, this also suggests that teaching gives added value.

For the English fluency scores, for the students there was a significant difference ($p<.01$) between semester one (mean score 26.5) and semester six (mean score 31.7). There was also a difference between semester six students and working interpreters (mean score 36.6, $p<.01$). This suggests that going to university professionalises your English but life experience working with English gives you greater fluency, which comes as no great surprise. There was also a promising correlation between English fluency and simultaneous interpreting exam results ($r=.557$, $p=0.07$). However, the small n makes it difficult to know whether this relationship would be strengthened with a larger sample size. It is also not clear whether this is because greater fluency allows for better interpreting performance or better English comprehension and production.

The perceptual processing (patterns) performance of the students in semester 6 were compared with scores of working interpreters using a t-test. There is no group difference between students and the working interpreters for perceptual processing (mean scores 43.1 vs 39.8, $p=0.161$). This suggests that although the student's BSL improves, this does not bring about a global improvement in perceptual speed. Any perceptual improvements are limited to perceiving BSL better and so is driven by a more robust phonological model and making sense of the language itself. This was confirmed by significantly higher scores in the non-sign repetition task in semester six when compared with semester one, although this data analysis is ongoing. It is worth noting that the non-sign repetition task generated a large amount of data that has proven difficult to code, which has impeded the data analysis. Currently data analysis is undertaken on an ad hoc basis by students interested in the phonological development of language learners and so this area of data analysis has been slow to move forwards.

Some t-tests were carried out to compare the performance of the semester 6 scores with working interpreters for psychomotor skills (connections A) and task-switching (connections B). There is no group difference between students and the working interpreters for psychomotor skills (mean scores 132.3 vs 132.1, $p=0.977$). There was a marginal group difference between BSL learners and expert interpreters for task switching (mean scores 63.4 vs 72.9, $p=0.091$).

For the flanker task an Independent samples t-test on the interference effect (response times, neutral minus incongruent) was undertaken. The students were tested three times and the working interpreters once. The difference between the students and the experts gradually reduced, such that there was only a significant difference in performance for semester three vs working interpreter ($p=0.010$) and semester five vs working interpreter ($p=0.036$) scores. By semester six these differences had washed out ($p=0.767$). This suggests that learning BSL as an adult increases some aspects of cognitive control more than others, e.g. conflict processing more than task switching.

In the main there were no differences detected between those rehoused in Deaf Studies and interpreting students. One result has proven difficult to interpret. When examining the semester five results of the flanker task for interference (response times, neutral minus incongruent) the interpreter student reaction time = -74ms vs Deaf Studies student reaction time = -39ms were significantly different ($p = .008$). It was unclear whether this indicated a window of learning for the interpreting students as they were gaining more control over the interpreting process.

Lastly the BIS did discriminate between the Deaf Studies students and Interpreting students in semester three. The interpreter students (mean score 69.9) were significantly different ($p = 0.017$) from the Deaf Studies students (mean score 59.8). This suggests that at least at that moment in the course the interpreting students were more willing to take risks than those who were rehoused (by choice or by exam performance) in Deaf Studies.

6. Conclusion

This exploratory longitudinal study, if anything, has confirmed that learning language, learning interpreting and interpreter expertise are complex phenomena. Much of the literature is inconclusive with respect to the variety of factors that contribute to language learning and interpreting, although it is clear that motivation is an important factor in language learning and this was not measured in this study. It would have been interesting to try and gauge the level of interaction that the students and the working interpreter had using BSL. This level of habitual usage of BSL, along with an account of the actual hours of interpreting that working interpreters felt/reported that they had undertaken could have provided further information, which would have enabled a regression analysis taking these factors into account.

Secondly, it is clear that although this study contributed to our understanding of some of the tests (BSLGJT and non-sign repetition task), it would have been more useful to have tests that had already been finalised to then use with this population. The contribution does however mean that, should these be used again in the UK, shorter versions can be used, which in turn could mean that other tests could be used alongside them, without making the test battery too long for participants to comfortably engage with. This could be of interest for other institutions that wish to screen their applicants for sign language interpreting programs.

Thirdly, there are some interesting factors that are emerging as contributing to sign language interpreter training. The language learning is clearly a bottleneck for those wishing to become interpreters as also noted by Gómez et al. (2007). Some of that is mitigated by general language learning abilities and having normal cognitive function. However, some improvements

in performance may be related to general improvement in students attending university. Some improvements may be due to increased use of both languages in ever increasing domains of use. And some of the improvements might be related to risk taking. This could support the need for students to gain fluency in sign language and a BA/BSc level education before training as a sign language interpreter. These general cognitive abilities could then be used, as they are in spoken language interpreter training, to bootstrap interpreting ability.

Finally, for a longitudinal study a larger cohort of students would be advisable so that statistical regressions can be used to understand which factors contribute the most. For this to happen would require at least ten years of further data collection as the number of students trained in university programs are small. With a new program having begun as a full time four-year university training program since the data collection for this study, collaborative working could enable robust data to be collected exploring the complex route taken by students to develop language fluency and interpreting proficiency.

References

- Bacon, S. M., & Finnemann, M. D. (1990). A study of the attitudes, motives, and strategies of university foreign language students and their disposition to authentic oral and written input. *The Modern Language Journal* 74, 459–473. doi:10.1111/j.1540-4781.1990.tb05338.x
- Baddeley, A.D., Emslie, H., & Nimmo-Smith, I. (1992). *The Speed and Capacity of Language Processing (SCOLP) Test*. Bury St Edmunds: Thames Valley Test Company.
- Brainmetric. (2004). Paced Auditory Serial Addition Test (PASAT). Retrieved from <http://www.brainmetric.com/products/pasat.htm>
- Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology*. Massachusetts: MIT Press.
- Brien, D., Brown, R., & Collins, J. (2002). *The organisation and provision of British Sign Language/English interpreters in England, Scotland and Wales*. London: HMSO
- Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training research and education* (pp.87-136). Pittsburgh: University of Pittsburgh Press. (ERIC Document Reproduction Service No. ED 038 051)
- Carroll, J. B., & Sapon, S. M. (2002). *Modern Language Aptitude Test*. Bethesda, MD: Second Language Testing.
- CFA. (2012). Interpreting: National occupational standards. Retrieved from <http://www.skillsfca.org/images/pdfs/National%20Occupational%20Standards/Languages%20and%20Intercultural%20Working/2006/Interpreting.pdf>
- Chen Pichler, D. & Koulidobrova, H., (2011). Acquisition of sign language as a second language (L2). In M. Marschark & P. E. Spencer, eds. *Oxford Handbook of Deaf Studies in Language*. Oxford University Press.
- Cormier, K., Schembri, A., Vinson, D., & Orfanidou, E. (2012). First language acquisition differs from second language acquisition in prelingually deaf signers: Evidence from sensitivity to grammaticality judgement in British Sign Language. *Cognition* 124(1), 50-56. <http://dx.doi.org/10.1016/j.cognition.2012.04.003> Retrieved from <http://www.sciencedirect.com/science/article/pii/S0010027712000753>
- Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition* 113 (2), 135–149. <http://dx.doi.org/10.1016/j.cognition.2009.08.001>

- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception & Performance* 14, 113-121.
- Emmorey, K., Luk, G., Pyers, J. E., & Bialystok, E. (2008a) The source of enhanced cognitive control in bilinguals: evidence from bimodal bilinguals. *Psychological Science* 19(12), 1201–1206. <http://doi.org/10.1111/j.1467-9280.2008.02224.x>
- Emmorey, K., Borinstein, H. B., Thompson, R., & Gollan, T. H. (2008b). Bimodal bilingualism. *Bilingualism* 11(1), 43–61.
- Ericsson, K.A. (2000/01). Expertise in interpreting: An expert-performance perspective. *Interpreting* 5(2), 187-220.
- Gómez, J. M. L., Molina, T. B., Benitez, P. P., & de Torres, J. S. (2007). Predicting proficiency in signed language interpreting: A preliminary study. *Interpreting* 9(1), 71-93.
- Gronwall, D.M. (1977). Paced auditory serial-addition task: a measure of recovery from concussion. *Percept Mot Skills* 44(2), 367–373
- Hedderly, R.G. (1992). Psychologists' assessments of specific learning difficulties (dyslexia) and examination boards: Policies and practices. *Educational Psychology In Practice* 8(1), 32-42.
- Liu, M., Schallert, D. L., & Carroll, P. J. (2004). Working memory and expertise in simultaneous interpreting. *Interpreting* 6 (1), 19-42.
- Macnamara, B., Adam, B.M., Kegl, J.A. & Conway A.R.A. (2011). Domain-general cognitive abilities and simultaneous interpreting skill. *Interpreting* 13(1), 121-142. doi 10.1075/intp.13.1.08mac
- Macnamara, B.N., & Conway, A.R.A., (2016). Working memory capacity as a predictor of simultaneous language interpreting performance. *Journal of Applied Research in Memory and Cognition*. Retrieved from www.sciencedirect.com/science/article/pii/S2211368115000923
- Mayberry, R., & Boudreault, P. (2006). Grammatical processing in American Sign Language: Age of first-language acquisition effects in relation to syntactic structure. *Language and Cognitive Processes* (5), 608-635
- NRCPD. (2016). NRCPD Newsletter September 2016. Retrieved from <http://www.nrcpd.org.uk/news.php?article=137>
- Office for Disability Issues HM Government. (2011). *UK Initial Report On the UN Convention on the Rights of Persons with Disabilities*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/345120/uk-initial-report.pdf
- Orfanidou, E., Adam, R., McQueen, J.M., & Morgan, G. (2009). Making sense of nonsense in British Sign Language (BSL): The contribution of different phonological parameters to sign recognition. *Memory & Cognition* 37(3), 302-315. doi:10.3758/MC.37.3.302
- Ortega, G., & Morgan, G. (2015). Phonological development in hearing learners of a sign language: The influence of phonological parameters, sign complexity, and iconicity. *Language Learning* 65, 660-688
- Patton, J.H., Stanford, M.S., & Barratt, E.S. (1995). Factor structure of the Barratt impulsivity scale. *Journal of Clinical Psychology* 51, 768-774.
- Pro-Sign. (2016) *Sign languages and the Common European Framework of Reference for Languages Common Reference Level Descriptors*. Retrieved from <http://www.ecml.at/Portals/1/mtp4/pro-sign/documents/Common-Reference-Level-Descriptors-EN.pdf>
- Russo, M. (2011). Aptitude testing over the years. *Interpreting* 13(1), 5–30. doi 10.1075/intp.13.1.02rus
- Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology* 27(5), 763-776.
- Salthouse, T. A., Toth, J., Daniels, K., Parks, C., Pak, R., Wolbrette, M., & Hocking, K.J. (2000). Effects of aging on efficiency of task switching in a variant of the trail making test. *Neuropsychology* 14(1), 102-111.
- Scott-Gibson, L. (1991). Sign Language Interpreting: An Emerging Profession. In S. Gregory and G. M. Hartley (Eds.) *Constructing Deafness* (pp. 253–58). London: Pinter in association with the Open University

- Simpson, S. (2007). *Advance to an ideal: The fight to raise the standard of communication between Deaf and hearing People*. Edinburgh: Scottish Workshop Publications.
- Sparks, R., Patton, J., Ganschow, L., & Humback, N. (2009). Long-term crosslinguistic transfer of skills from L1 to L2. *Language Learning* 59(1), 203-243.
- Stone, C. (2013). The UNCRPD and 'professional' sign language interpreter provision. In C. Schäffner, K. Kredens and Y. Fowler (Eds.) *Interpreting in a Changing Landscape. Selected papers from Critical Link 6* (pp. 83-100). Amsterdam: John Benjamins.
- Stone, C. & Woll, B. (2008). DUMB O JEMMY and others: Deaf people, interpreters and the London courts in the 18th and 19th centuries, *Sign Language Studies* 8 (3), 226 - 240. doi: 10.1353/sls.2008.0009
- Timarová, S., Čeňková, I., Meylaerts, R., Hertog, E., Szmalec, A., & Duyck, W. (2014). Simultaneous interpreting and working memory executive control. *Interpreting* 16(2), 139–168. doi 10.1075/intp.16.2.01tim
- Wang, J., & Napier, J. (2013). Signed language working memory capacity of signed language interpreters and deaf signers. *Journal of Deaf Studies and Deaf Education* 18 (2), 271-286. doi:10.1093/deafed/ens068
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. New York: The Psychological Corporation: Harcourt Brace & Company.